

Lanne, Markku; Lütkepohl, Helmut; Saikkonen, Pentti

Working Paper

Test procedures for unit roots in time series with level shifts at unknown time

SFB 373 Discussion Paper, No. 2001,39

Provided in Cooperation with:

Collaborative Research Center 373: Quantification and Simulation of Economic Processes, Humboldt University Berlin

Suggested Citation: Lanne, Markku; Lütkepohl, Helmut; Saikkonen, Pentti (2001) : Test procedures for unit roots in time series with level shifts at unknown time, SFB 373 Discussion Paper, No. 2001,39, Humboldt University of Berlin, Interdisciplinary Research Project 373: Quantification and Simulation of Economic Processes, Berlin, <https://nbn-resolving.de/urn:nbn:de:kobv:11-10049801>

This Version is available at:

<https://hdl.handle.net/10419/62711>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Test Procedures for Unit Roots in Time Series with Level Shifts at Unknown Time

by

Markku Lanne

University of Helsinki

Helmut Lütkepohl

Humboldt University Berlin

and

Pentti Saikkonen

University of Helsinki

Address for correspondence: Markku Lanne, Department of Economics, PO Box 54, FIN-00014 University of Helsinki, FINLAND

Abstract

Two types of unit root tests which accommodate a structural level shift at a known point in time are extended to the situation where the break date is unknown. It is shown that for any estimator for the break date the tests have the same asymptotic distribution as the corresponding tests under the known break date assumption. Different estimators of the break date are compared in a Monte Carlo experiment and a recommendation for choosing the break date in small samples is given. It is also shown that ignoring the fact that a break has occurred and applying a standard unit root test may lead to substantial size distortion and total loss of power. Example series from the Nelson-Plosser data set are used to illustrate the performance of our tests.

JEL classification: C22, C12

Keywords: Univariate time series, unit root, structural shift, autoregression

*We are grateful to Stefan Fleck for helping with the computations and to Christian Müller for useful comments on an earlier version of the paper. Moreover, we thank the Deutsche Forschungsgemeinschaft, SFB 373, the European Commission under the Training and Mobility of Researchers Programme (contract No. ERBFMRXCT980213) and the Yrjö Jahnsson Foundation for financial support. The third author also thanks the Alexander von Humboldt Foundation for financial support under a Humboldt research award. Part of this research was carried out while the first and third authors were visiting the Humboldt University in Berlin.

1 Introduction

Unit root tests are commonly used at the beginning of an econometric time series analysis. Therefore suitable tests for different situations have been proposed in the literature. In particular, a number of articles consider unit root tests in the presence of possible structural breaks. In the related literature two alternative assumptions regarding the possible dates of the structural breaks have been made. Some authors assume that the break date is known a priori whereas others assume that the break date is endogenous and is therefore unknown to the analyst. The known break date assumption has been made by Perron (1989, 1990), Saikkonen & Lütkepohl (2001a), Lütkepohl, Müller & Saikkonen (2001) and Lanne, Lütkepohl & Saikkonen (2001) (henceforth LLS) among others. An unknown break date is assumed, for example, by Evans (1989), Christiano (1992), Perron & Vogelsang (1992), Zivot & Andrews (1992), Banerjee, Lumsdaine & Stock (1992), Leybourne, Newbold & Vougas (1998) and Saikkonen & Lütkepohl (2001b) (henceforth S&L).

If the break date is unknown, it has to be estimated from the given time series. In this study we will consider unit root tests in the presence of a level shift and compare the impact of different proposals for specifying the break date on the properties of the tests. We will do so in the context of unit root tests proposed by LLS for the case where the break date is known. These tests will be extended to the unknown break date case. The reason is that LLS show that their tests have favourable properties if the break date is known. Therefore, we expect them to work well in the unknown break date case as well. Furthermore, the tests have the advantage that they allow for very general shift functions and not just an abrupt shift as assumed in some of the related literature. The model considered by LLS is sometimes referred to as additive outlier model in the literature. We will also consider innovational outlier models and present the corresponding unit root tests. An important advantage of the tests proposed in the following is that the asymptotic distribution does not depend on the break date or the estimator of the break date.

In the literature on unit root tests for series with a structural shift at unknown time, different estimators for the break date have been proposed. Keeping in mind the objective of testing for a unit root, the break date is chosen in such a way that the considered tests apply. Perron & Vogelsang (1992) and Zivot & Andrews (1992) propose to estimate the break date such that the unit root test becomes least favourable to the null hypothesis of a

unit root and they develop the asymptotic distribution theory of the resulting test statistic. Banerjee, Lumsdaine & Stock (1992) consider the asymptotic distributions of unit root tests based on recursively computed test statistics. In contrast, Leybourne, Newbold & Vougas (1998) estimate the deterministic part of the assumed data generation process (DGP) first, including possible structural shifts and then they apply unit root tests to the residuals. In the present study we will compare different approaches for choosing the break date in a Monte Carlo study. The DGP is assumed to have an autoregressive (AR) structure. Therefore, we will also discuss strategies for choosing the AR order.

The structure of the paper is as follows. In the next section two general models for univariate time series with a shift in the mean and a possible unit root are presented. The models are those treated by S&L. Section 3 considers estimation of the nuisance parameters of the DGP and the tests for unit roots and estimators for the break date are presented in Section 4. In Section 5 a Monte Carlo comparison of different tests and strategies for choosing the break date is discussed. Examples are considered in Section 6 and conclusions are contained in Section 7. Proofs are given in the Appendix.

The following general notation is used. The lag and differencing operators are denoted by L and Δ , respectively, so that for a time series variable y_t , $Ly_t = y_{t-1}$ and $\Delta y_t = y_t - y_{t-1}$. The symbol \xrightarrow{d} signifies convergence in distribution. Independently, identically distributed will be abbreviated as $iid(\cdot, \cdot)$, where the first and second moments are indicated in parentheses in the usual way. Furthermore, $O(\cdot)$, $o(\cdot)$, $O_p(\cdot)$ and $o_p(\cdot)$ are the usual symbols for the order of convergence and convergence in probability, respectively, of a sequence. We use $\lambda_{\min}(A)$ to denote the minimal eigenvalue of a matrix A . Moreover, $\|\cdot\|$ denotes the Euclidean norm. LS is used to abbreviate least squares and sup and inf are short for supremum and infimum, respectively. The m -dimensional Euclidean space is denoted by \mathbf{R}^m .

2 DGPs

Two alternative models for the DGP of a time series with a possible unit root and a level shift are considered. The first one is

$$y_t = \mu_0 + \mu t + f_{t\eta}(\theta)' \gamma + x_t, \quad t = 1, 2, \dots, \quad (2.1)$$

where the scalar μ , the $(m \times 1)$ vector θ and the $(k \times 1)$ vector γ are unknown parameters and $f_{t\eta}(\theta)$ is a $(k \times 1)$ vector of deterministic sequences depending on the parameters θ and on the break point which is denoted by η , that is, a shift occurs in or just before period $\eta \in \{1, \dots, T\}$. The stochastic process x_t is assumed to have an AR representation of order p ,

$$b(L)(1 - \rho L)x_t = \varepsilon_t, \quad (2.2)$$

where $b(L) = 1 - b_1L - \dots - b_{p-1}L^{p-1}$ has all its zeros outside the unit circle if $p > 1$, while $-1 < \rho \leq 1$. A unit root is present if $\rho = 1$. The initial values x_t , $t \leq 0$, are assumed to be independent of the sample size T . The error terms ε_t are assumed to be *iid*(0, σ^2) with $E|\varepsilon_t|^\alpha < \infty$ for some $\alpha > 4$.

The second model has the form

$$b(L)y_t = \mu_0 + \mu t + f_{t\eta}(\theta)' \gamma + v_t, \quad t = 1, 2, \dots, \quad (2.3)$$

where

$$v_t = \rho v_{t-1} + \varepsilon_t \quad (2.4)$$

is an AR process of order 1 and the other notation is as before. Again, if $\rho = 1$, v_t and, hence, y_t has a unit root. The presample values y_{-p+1}, \dots, y_0 in (2.3) are assumed to be observable in addition to the sample values.

The sequence $f_{t\eta}(\theta)$ is assumed to satisfy the following conditions adapted from LLS for the presently considered case of an unknown break date.

Assumption A.

- (a) The parameter space of θ , denoted by Θ , is a compact subset of \mathbf{R}^m and N_T , the space of η , is a subset of $\{2, \dots, T - 1\}$.
- (b) For each $t = 1, 2, \dots$ and each $\eta \in N_T$, $f_{t\eta}(\theta)$ is a continuously differentiable function in an open set containing the parameter space Θ and, denoting by $F_{t\eta}(\theta)$ the vector of all partial derivatives of $f_{t\eta}(\theta)$ with respect to θ ,

$$\sup_T \sup_{\theta \in \Theta, \eta \in N_T} \sum_{t=1}^T \|\Delta f_{t\eta}(\theta)\| < \infty \quad \text{and} \quad \sup_T \sup_{\theta \in \Theta, \eta \in N_T} \sum_{t=1}^T \|\Delta F_{t\eta}(\theta)\| < \infty$$

where $f_{0\eta}(\theta) = 0$ and $F_{0\eta}(\theta) = 0$.

(c) $f_{1\eta}(\theta) = \dots = f_{p+1,\eta}(\theta) = 0$ for all $\theta \in \Theta$ and all $\eta \in N_T$. Moreover, defining $G_{t\eta}(\theta) = [f_{t\eta}(\theta)' : F_{t\eta}(\theta)']'$ for $t = 0, 1, 2, \dots$, there exists a real number $\epsilon > 0$ and an integer T_* such that, for all $T \geq T_*$,

$$\inf_{\theta \in \Theta, \eta \in N_T} \lambda_{min} \left\{ \sum_{t=1}^T \Delta G_{t\eta}(\theta) \Delta G_{t\eta}(\theta)' \right\} \geq \epsilon.$$

□

In Assumption A(a) it is specified that the break date cannot be at the very beginning or at the very end of the sample but it may be confined to a subset of the remaining sample if, e.g., prior knowledge about the period exists where it may have occurred. The other assumptions are related to the types of shift functions that may be considered. Roughly speaking, abrupt shifts as well as smooth shifts from one level to another are permitted. For instance, for $\eta > p + 1$, possible functions $f_{t\eta}(\theta)$ satisfying Assumption A are

$$f_{t\eta}(\theta) = d_{t\eta} := \begin{cases} 0, & t < \eta \\ 1, & t \geq \eta \end{cases}, \quad (2.5)$$

$$f_{t\eta}(\theta) = \begin{cases} 0, & t < \eta \\ 1 - \exp\{-\theta(t - \eta + 1)\}, & t \geq \eta \end{cases} \quad (2.6)$$

or

$$f_{t\eta}(\theta) = \begin{cases} 0, & t < \eta \\ \exp\{-\theta(t - \eta)\}, & t \geq \eta \end{cases},$$

where θ is an unknown parameter with $0 < \theta < \text{constant} < \infty$, and

$$f_{t\eta}(\theta) = \left[\frac{d_{t,\eta}}{\varphi(L)} : \dots : \frac{d_{t-q,\eta}}{\varphi(L)} \right]',$$

where the components of θ are given by the unknown coefficients of $\varphi(L) = 1 - \varphi_1 L - \dots - \varphi_r L^r$, which is a lag polynomial with all its zeros outside the complex unit circle.

In Assumption A(c) the condition $f_{1\eta}(\theta) = \dots = f_{p+1,\eta}(\theta) = 0$ for $\eta \in N_T$ is satisfied for these functions if η is known to be greater than $p + 1$ (i.e., $N_T \subset \{p + 2, \dots, T - 1\}$). This condition together with the last condition in Assumption A(c) implies that

$$\inf_{\theta \in \Theta, \eta \in N_T} \lambda_{min} \left\{ \sum_{t=p+2}^T [b(L) \Delta G_{t\eta}(\theta)] [b(L) \Delta G_{t\eta}(\theta)'] \right\} \geq \epsilon$$

for $T \geq T_*$ which is needed for some of the estimators used in the following to be well-defined.

Our unit root tests proceed by estimating the parameters of the deterministic part first and subtracting it from the series. Then standard unit root tests are applied to the adjusted series. In the next section, the estimation procedures for the parameters of the deterministic part from S&L are reviewed and the unit root tests are discussed in Section 4.

3 Estimation of Nuisance Parameters

We begin by discussing the estimation of the nuisance parameters in model (2.1)/(2.2). Applying the filter $\Delta = 1 - L$ yields the model

$$Y = Z_\eta(\theta)\phi + U, \quad (3.1)$$

where $Y = [y_1 : \Delta y_2 : \dots : \Delta y_T]'$, $\phi = [\mu : \mu_0 : \gamma]'$, $Z_\eta(\theta) = [Z_1 : Z_{2\eta}(\theta)]$ with $Z_1 = [1 : \dots : 1]'$ and $Z_{2\eta}(\theta) = [g_{1\eta}(\theta) : \Delta g_{2\eta}(\theta) : \dots : \Delta g_{T\eta}(\theta)]'$ with $g_{t\eta}(\theta) = [1 : f_{t\eta}(\theta)]$. Finally, $U = [u_1 : \dots : u_T]'$ is an error term such that $u_t = x_t - x_{t-1} = b(L)^{-1}\varepsilon_t$ if a unit root exists.

Under this assumption, for any given values of η and p the parameters θ and ϕ as well as the parameters $b = [b_1, \dots, b_{p-1}]'$ in the error covariance matrix of U can be estimated by minimizing the generalized sum of squares function

$$Q_{\eta,p}(\phi, \theta, b) = (Y - Z_\eta(\theta)\phi)' \Sigma_p(b)^{-1} (Y - Z_\eta(\theta)\phi), \quad (3.2)$$

where $\Sigma_p(b) = \sigma^{-2} \text{Cov}(U)$. The properties of the resulting estimators are given in S&L.

For estimating the parameters of the model (2.3) we also premultiply by Δ and get

$$Y = W_{\eta,p}(\theta)\beta + \mathcal{E}, \quad (3.3)$$

where $\beta = [b' : \phi']'$, $W_{\eta,p}(\theta) = [V_p : Z_\eta(\theta)]$ with V_p the $(T \times (p-1))$ matrix containing lagged values of the regressand. Furthermore, $\mathcal{E} = [\varepsilon_1 : \dots : \varepsilon_T]'$ if the unit root hypothesis is satisfied. For given η and p , the estimators are obtained by minimizing

$$S_{\eta,p}(\theta, \beta) = (Y - W_{\eta,p}(\theta)\beta)' (Y - W_{\eta,p}(\theta)\beta). \quad (3.4)$$

The asymptotic properties of the resulting estimators are also given in S&L.

An estimator of η is, of course, needed to make the estimators of the nuisance parameters feasible. If $\hat{\eta}$ is some estimator of η , feasible counterparts of the nuisance parameters are

defined in an obvious way. It turns out that the asymptotic properties of the unit root tests to be studied in the next section do not depend on the choice of the estimator $\hat{\eta}$. However, the choice of $\hat{\eta}$ may be important in finite samples and this is what we intend to explore in Section 5.

4 Unit Root Tests

4.1 Tests for Given p and $\hat{\eta}$

For given AR order p and shift date estimate $\hat{\eta}$ we can estimate the nuisance parameters as described in the previous section and obtain adjusted series on which the unit root tests may be based. For model (2.1) we get $\hat{x}_t = y_t - \hat{\mu}_{0\hat{\eta}} - \hat{\mu}_{\hat{\eta}}t - f_{t\hat{\eta}}(\hat{\theta}_{\hat{\eta}})'\hat{\gamma}_{\hat{\eta}}$ and for model (2.3) $\tilde{v}_t = \tilde{b}_{\hat{\eta}}(L)y_t - \tilde{\mu}_{0\hat{\eta}} - \tilde{\mu}_{\hat{\eta}}t - f_{t\hat{\eta}}(\tilde{\theta}_{\hat{\eta}})'\tilde{\gamma}_{\hat{\eta}}$. There are several possible unit root tests that can be used. We focus on versions which performed well in a simulation comparison by LLS. In these tests adjustments for the estimation uncertainty in the nuisance parameters are made.

Define $w_t = b(L)x_t$ so that $w_t = \rho w_{t-1} + \varepsilon_t$. Thus, if we condition on y_1, \dots, y_p , a version of the test statistic in the context of model (2.1) may be based on the auxiliary regression model

$$\hat{w}_t = \rho \hat{w}_{t-1} + error_t, \quad t = p+1, \dots, T,$$

where $\hat{w}_t = \hat{b}_{\hat{\eta}}(L)\hat{x}_t$. To adjust for the estimation errors in the nuisance parameters LLS find that a test based on the auxiliary regression model

$$\hat{w}_t = \nu + \rho \hat{w}_{t-1} + [\hat{b}_{\hat{\eta}}(L)\Delta f_{t\hat{\eta}}(\hat{\theta}_{\hat{\eta}})']\pi_1 + [\hat{b}_{\hat{\eta}}(L)\Delta F_{t\hat{\eta}}(\hat{\theta}_{\hat{\eta}})']\pi_2 + \sum_{j=1}^p \alpha_j \Delta \hat{x}_{t-j} + r_t, \quad t = p+2, \dots, T, \quad (4.1)$$

is preferable. The unit root test statistic is obtained as the usual t -statistic for the hypothesis $\rho = 1$ based on LS estimation of this model. It will be denoted by $\tau(\hat{\eta}, p)$ if a linear trend term is included and by $\tau^0(\hat{\eta}, p)$ if $\mu = 0$ a priori.

To formulate a test based on the model (2.3) we consider mean-adjusted variables $\Delta y_{t-j} - \tilde{\mu}_{*\hat{\eta}}$ ($j = 1, \dots, p$) where $\tilde{\mu}_{*\hat{\eta}} = \tilde{\mu}_{\hat{\eta}}/\tilde{b}_{\hat{\eta}}(1)$. Defining $\tilde{q}'_{t\hat{\eta}} = [\Delta y_{t-1} - \tilde{\mu}_{*\hat{\eta}} : \dots : \Delta y_{t-p+1} - \tilde{\mu}_{*\hat{\eta}}]$, we consider the auxiliary regression model

$$\tilde{v}_t = \nu + \rho \tilde{v}_{t-1} + \Delta f_{t\hat{\eta}}(\tilde{\theta}_{\hat{\eta}})'\pi_1 + \Delta F_{t\hat{\eta}}(\tilde{\theta}_{\hat{\eta}})'\pi_2 + \tilde{q}'_{t\hat{\eta}}\pi_3 + e_t^\dagger, \quad t = 2, \dots, T, \quad (4.2)$$

and propose to use the relevant t -statistic denoted by $\mathbf{t}(\hat{\eta}, p)$ for a unit root test. The corresponding statistic based on a model without time trend ($\mu = 0$) will be denoted by $\mathbf{t}^0(\hat{\eta}, p)$. In the next theorem the asymptotic distributions under local alternatives $\rho = 1 + \frac{c}{T}$, $c \leq 0$, are given for an arbitrary estimator $\hat{\eta}$ of the break point η .

Theorem 1.

Suppose that Assumption A stated in Section 2 holds and assume that, for some $\epsilon > 0$, $b(L) \neq 0$ for $|L| \leq 1 + \epsilon$, that is, the roots of $b(L)$ are bounded away from the unit circle. Moreover, suppose that the matrix $Z_\eta(\theta)$ is of full column rank for all $\theta \in \Theta$, all $\eta \in N_T$ and all $T \geq k + 2$. Then, denoting by $B_0(u)$ a standard Brownian motion and using the Ornstein-Uhlenbeck process $B_c(s) = \int_0^s \exp\{c(s-u)\}dB_0(u)$ and its mean-adjusted counterpart $\bar{B}_c(s) = B_c(s) - \int_0^1 B_c(u)du$,

$$\tau^0(\hat{\eta}, p), \mathbf{t}^0(\hat{\eta}, p) \xrightarrow{d} \left(\int_0^1 \bar{B}_c(s)^2 ds \right)^{-1/2} \int_0^1 \bar{B}_c(s)dB_c(s). \quad (4.3)$$

Moreover,

$$\tau(\hat{\eta}, p), \mathbf{t}(\hat{\eta}, p) \xrightarrow{d} \left(\int_0^1 \bar{G}_c(s)^2 ds \right)^{-1/2} \int_0^1 \bar{G}_c(s)dG_c(s), \quad (4.4)$$

where $\bar{G}_c(s)$ is a mean-adjusted version of $G_c(s) = B_c(s) - sB_c(1)$. □

The asymptotic distributions in the theorem are precisely the same as for the corresponding tests based on series with known break date (see LLS for the asymptotic distributions of the τ tests when the break date is known). Therefore the critical values from Table 2 of LLS may be used in the tests. Note that LLS also consider estimation of the nuisance parameters under local alternatives which results in more general versions of the asymptotic distributions. Because LLS found that estimation under local alternatives does not improve the small sample properties of the tests, we do not consider these generalizations here. LLS also present a number of other tests which we do not discuss here because they were found to have inferior small sample properties in the case of a known break date. As mentioned previously, the τ tests presented here are those which performed best in the additive outlier model of LLS.

In practice, the AR order p and the break date estimate $\hat{\eta}$ have to be chosen in such a way that the actual small sample distributions of our test statistics resemble the asymptotic

null distributions as closely as possible and best possible power is obtained. Alternative strategies for choosing p and $\hat{\eta}$ are presented in the following subsection.

4.2 Specifying p and Estimating η

There are a number of different ways for choosing p and η . The choice of the AR order used in estimating the nuisance parameters may in principle differ from the choice used in the unit root tests in (4.1) and (4.2). In this section we will present some strategies which will be compared in the simulation section. It will be seen that the choice of p is not critical for the unit root tests as long as it is chosen reasonably large. We will therefore focus on the estimation of the break date first and then briefly comment on AR order selection.

The following estimators for η will be considered for some given AR order p :

- $\hat{\eta}(p) = \arg \min_{\eta \in N_T} Q_{\eta,p}(\phi, \theta, b)$ or $\hat{\eta}(p) = \arg \min_{\eta \in N_T} S_{\eta,p}(\phi, \beta)$, depending on which model is used.
- $\tilde{\eta}(p) = \arg \min_{\eta \in N_T} \tau(\eta, p)$ [or $\tilde{\eta}(p) = \arg \min_{\eta \in N_T} \tau^0(\eta, p)$] or $\tilde{\eta}(p) = \arg \min_{\eta \in N_T} \mathbf{t}(\eta, p)$ [or $\tilde{\eta}(p) = \arg \min_{\eta \in N_T} \mathbf{t}^0(\eta, p)$], depending on which model is used. The idea here is to choose the shift date such that it is least favorable for the unit root null hypothesis.

Both selection strategies require the estimation of the nuisance parameters for many values of η if N_T is a large set. The first estimator is equivalent to choosing the break date such that the F -statistic for testing the null hypothesis $\gamma = 0$ is maximized. The latter estimation method was considered by Perron & Vogelsang (1992).

In estimating η , a reasonably large p should be chosen to ensure that the residuals of the model with break date $\hat{\eta}(p)$ or $\tilde{\eta}(p)$ are white. Choosing the order slightly too large does not seem to be critical (see the simulations in the next section). Alternatively one may determine $\hat{\eta}(p)$ or $\tilde{\eta}(p)$ for $p = 1, \dots, p_{\max}$ and choose p such that it minimizes a model selection criterion

$$C_Q(p) = \log Q_{\hat{\eta},p}(\hat{\phi}, \hat{\theta}, \hat{b}) + c_T p/T$$

or

$$C_S(p) = \log S_{\hat{\eta},p}(\hat{\phi}, \hat{\beta}) + c_T p/T$$

depending on the model used. Here $c_T = 2, 2 \log \log T$ or $\log T$ for the well-known model selection criteria AIC, HQ or SC, respectively (see, e.g., Lütkepohl (1991) for a more detailed

discussion of these model selection criteria). Yet another possibility is to determine $\hat{\eta}(p_{\max})$ or $\tilde{\eta}(p_{\max})$ and reduce the order p only if the corresponding unit root test cannot reject the unit root null hypothesis and the desired model selection criterion permits a reduction. This strategy may be helpful in reducing the computational burden.

5 Monte Carlo Study

We have compared the unit root tests for different estimators of the break date and we have investigated the impact of varying the AR order via a simulation experiment based on the following two processes:

$$y_t = \gamma d_{t\eta} + x_t, \quad (1 - b_1 L)(1 - \rho L)x_t = \varepsilon_t, \quad t = 1, \dots, T, \quad (5.1)$$

and

$$(1 - b_1 L)y_t = \gamma d_{t\eta} + v_t, \quad v_t = \rho v_{t-1} + \varepsilon_t, \quad t = 1, \dots, T, \quad (5.2)$$

with $\varepsilon_t \sim iid N(0, 1)$, $\rho = 1, 0.9, 0.8$, $T = 100, 200$ and different values of γ . Here $d_{t\eta}$ is the shift dummy variable defined in (2.5). The processes are similar to those used in a Monte Carlo study by LLS. In the simulations we also generated 100 presample values which were discarded except that presample values were used in the estimations underlying model (2.3). The first process (5.1) is a special case of the model (2.1) with an abrupt shift at time η so that the τ tests are the appropriate tests whereas in general the model underlying the \mathbf{t} tests can only approximate the DGP (5.1). Thus applying this test as well should give some indication of the flexibility of the framework and of the consequences of using a misspecified model. In turn, the DGP (5.2) is a special case of (2.3). For $b_1 \neq 0$, it generates a smooth shift in the deterministic term. The \mathbf{t} tests are designed for this process, whereas the τ tests are approximations only. To capture the smooth transition from one regime to another the τ tests may be combined with a smooth shift function. For both types of tests we use different shift functions from the collection given in Section 2. Although there is no linear trend term in the DGPs we allow for such a term in computing some of the test statistics. In other words, in some of our simulations we assume that the absence of a linear trend is not known a priori. The more important results from our simulations are presented in a series of figures.

In Figure 1, the frequency distributions of estimated break dates are shown in histograms with 15 bins of equal length, using the true AR order $p = 2$ and assuming that the type of DGP including the shift function is known. Thus, we are operating under ideal conditions. Notice, however, that we have considered every fifth observation only as a possible shift date in order to economize on computer time. Obviously, $\hat{\eta}$ finds the true break date of $\eta = 50$ more often than $\tilde{\eta}$. Note, however, that for small shifts ($\gamma = 1$) both estimators have problems finding the shift date whereas for $\gamma = 3$, $\hat{\eta}$ locates the shift date with a very high probability. Similar results were also obtained for other shift dates and are not shown to save space. Specifically, we have also performed simulations for $\eta/T = 0.1, 0.3, 0.7$ and 0.9 and found that the relative performance of the break date estimators is similar to that shown in Figure 1. Of course, strictly speaking the frequency of correct choices of the break date may not be the best performance criterion here because we are eventually interested in the implications for the unit root tests. It turns out, however, that the estimator $\hat{\eta}$ also results in the best performance of the unit root tests. Therefore, we will focus on $\hat{\eta}$ in the following.

In Figure 2, power functions for different break dates η are displayed. Again we are working under ideal conditions because we assume that the type of DGP and the type of shift are given. In other words, the shift function is represented by a simple shift dummy variable. Also the true AR order $p = 2$ is used. In addition to the power functions of our τ and \mathbf{t} tests we also show those of standard augmented Dickey-Fuller (DF) tests which do not take the shift into account. It turns out that for the DGP (5.1) our tests are not sensitive to the shift date whereas the size and power of the DF tests clearly depends on the location of the shift. In the figure we only show results for tests without time trend. The results for our tests with time trend were similar whereas the DF tests with a linear trend are somewhat less sensitive to variations in the break date than the DF tests without time trend. In this respect the sensitivity of our tests as well as that of the DF tests is greater for the DGP (5.2) as can also be seen in Figure 2. Obviously, the DF tests may suffer dramatic size distortions and total loss of power for this DGP. From a practical point of view the tests become useless. In the following analyses we place the break date in the middle of the sample because this choice is the worst case for our tests in terms of power.

The impact of the size of the break on the power of the tests is shown in Figure 3. Not

surprisingly, the tests perform best if the DGP is the one assumed in their derivation. In other words, the τ tests may lose their power if (5.2) is the DGP and the \mathbf{t} tests tend to perform poorly if (5.1) is the actual DGP. The deterioration in power increases with the size of the shift. Another feature of the tests is also seen in Figure 3. It turns out that a priori knowledge of the absence of a linear trend is helpful for improving the power for DGP (5.1). In other words, for this DGP τ^0 and \mathbf{t}^0 have larger power than τ and \mathbf{t} , respectively. On the other hand, somewhat surprisingly, a corresponding power advantage is not obtained for DGP (5.2).

Because the shift generated by (5.2) is actually a smooth one it is tempting to expect the performance of the τ tests to improve if they are used together with a smooth shift function. Therefore we have applied both types of tests with a shift function (2.6) which produces an abrupt shift if θ is very large and may also capture smooth shifts in the mean for smaller values of θ . Some results are shown in Figure 4. Here θ is estimated along with the other parameters of the deterministic part. Obviously, in this case the performance of the tests deteriorates substantially. Both size and power are affected. In particular, the actual rejection frequencies of the \mathbf{t} tests exceed the nominal 5% substantially if there is actually a unit root. Moreover, the power of the τ tests is quite low in all the situations shown in Figure 4 in comparison to the corresponding cases in Figure 3. In other words, the additional flexibility in the shift function has a negative impact on the properties of the tests. It may be worth noting, however, that the performance of the tests with flexible shift functions improved markedly for larger sample sizes. Still, with samples of the size $T = 100$ it may be preferable to use the tests in conjunction with a simple shift dummy if the shift date is unknown. Given that the relative performance of our tests strongly depends on the type of DGP, applying both tests simultaneously may be a good strategy in practice where the true DGP is unknown. A unit root is then rejected if one of the tests rejects the null hypothesis because that test is regarded as the more powerful one.

Finally, in Figure 5, the impact of the AR order on the properties of the tests is illustrated. Clearly, under the ideal conditions underlying the figure (type of DGP and shift function known, actual AR order finite), overstating the AR order a bit reduces the power of the tests slightly. A much stronger effect may be induced by understating the order (see $p = 1$ and DGP (5.1)). In that case the power in particular of the τ and DF tests may drop

substantially. Therefore, it may be a good strategy to choose an AR order which ensures white residuals in the fitted model. For example, in a first step one may choose a fairly large order which is regarded as reasonable for a given time series. If the unit root hypothesis is rejected, no further adjustments are necessary. On the other hand, if the unit root is not rejected, it may be worth checking if a rejection is obtained with a smaller AR order.

6 Examples

We use annual U.S. time series from the well-known Nelson & Plosser (1982) data set as extended by Kleibergen & Hoek (1999) to illustrate the performance of the tests in practice.¹ Similar data have been analyzed by a number of other authors as well. A comparison with some previous results will be given later.

We have chosen ten series which start in the 19th century or in the early part of the 20th century and terminate in 1988. The precise sample periods are given in Table 1 together with the results of unit root tests. In addition to the sample values there are presample values for each of the series which are used in the estimation. All series are in natural logarithms and they are such that including a linear time trend seems justified.

In Table 1 results of DF, τ and \mathbf{t} tests are given. The lag order is chosen by estimating models up to order $p - 1 = 4$ with a linear trend and without break and we have then considered the t -ratios of the coefficients associated with the largest lags. The greatest order is selected for which the coefficient of the largest lag is significant at conventional levels. In some cases we also give results for models with larger lag orders if there is a specific reason for that. A discussion will be provided later. According to our simulation results, the choice of the AR order is not crucial and therefore we believe that the present approach is justified. For the given AR orders the break dates for the τ and \mathbf{t} tests are estimated using the estimator $\hat{\eta}$ in conjunction with a shift function consisting of a simple shift dummy variable as suggested by our simulation results.

The DF tests reject the unit root null hypothesis for four of the ten series at a 10% level of significance. In contrast, the τ tests reject for six of the series. In fact, they reject for all the series for which the DF tests also reject and in addition they indicate stationarity of the

¹We thank Frank Kleibergen for providing the data.

Table 1. Unit Root Tests for Annual U.S. Data

Variable	Sample period	$p - 1$	DF test ^a	τ test ^b		t test ^b	
				statistic	$\hat{\eta}$	statistic	$\hat{\eta}$
Real GNP	1914 - 1988	1	-3.60**	-2.89*	1933	-2.24	1923
		2	-3.48**	-2.72	1933	-2.67	1933
Nominal GNP	1914 - 1988	1	-2.00	-2.56	1922	-2.18	1922
		2	-1.73	-2.29	1922	-1.91	1922
Real per capita GNP	1914 - 1988	1	-3.68**	-2.89*	1933	-2.32	1923
		2	-3.57**	-2.73*	1933	-2.71*	1933
Industrial production	1865 - 1988	0	-3.42**	-3.56**	1922	-3.50**	1922
		1	-3.71**	-4.28**	1922	-3.91**	1922
Employment	1895 - 1988	1	-4.10**	-3.11**	1947	-2.93*	1947
GNP deflator	1894 - 1988	1	-1.58	-1.89	1922	-2.31	1922
Consumer prices	1865 - 1988	2	-1.60	-1.98	1921	-1.47	1922
		3	-1.89	-3.37**	1921	-1.67	1922
Nominal wage	1905 - 1988	1	-2.25	-2.60	1909	-2.82*	1922
Money stock	1894 - 1988	1	-2.91	-3.06**	1922	-2.96*	1922
		3	-2.82	-2.73	1922	-3.27**	1933
Velocity	1874 - 1988	0	-1.57	-1.39	1882	-1.56	1882
		1	-1.72	-1.46	1882	-1.82	1882

^aCritical values: -3.41 (5%), -3.12 (10%) (see Fuller (1976), Table 8.5.2, $\hat{\tau}_\tau$, $n = \infty$).
^bCritical values: -3.03 (5%), -2.76 (10%) (see Lanne et al. (2001), Table 2, $T = 1000$).
 ‘*’ and ‘**’ indicate significance at the 10% and 5% level, respectively.

consumer price index (CPI) and the money stock. This result is in line with our simulations which indicate that the DF tests are unreliable if there is a break and in particular they may have low power. It may be worth noting, however, that the τ tests reject the unit root for the CPI and the money stock only for one of the two AR orders given in Table 1. Although the third lag has a t -value of around 1.5 and is hence not significant at conventional levels, including it results in a substantial increase in the τ statistic. Given our simulation results, such a behaviour would be expected if the order $p - 1 = 2$ is in fact too small to capture the short-run dynamics adequately.

For the money stock variable, the situation is just the other way round. In this case the unit root is rejected by the τ test for the smaller lag order $p - 1 = 1$. Notice, however, that

the test value for $p - 1 = 3$ is very close to the critical value for a 10% significance level. Thus, in this case there may be a power decline due to the increase in the order. Both orders are given although the lag three coefficient is not significant at the 10% level because varying the order has a sizeable impact on the \mathbf{t} test. Not only does it reject the unit root with more confidence when the AR order is increased but it also picks a quite different shift date in this case. For order $p - 1 = 1$ the shift year is estimated to be 1922 whereas for $p - 1 = 3$ an estimate $\hat{\eta} = 1933$ is obtained. The result may be an indication that the shift is not very severe or that the model underlying the \mathbf{t} test is not a good representation of the DGP. Of course, it is also possible that there is more than one shift. Because we can reject the unit root even when only one shift is allowed for, there is no reason for a further exploration of that possibility.

Generally, the \mathbf{t} tests reject the null hypothesis less often than the τ tests in our data set. Given the power results obtained from the simulations in the previous section this outcome is clearly not surprising. It may be a bit surprising, however, that there is even one series for which the DF test rejects whereas \mathbf{t} does not at a 10% level of significance. This occurs for real gross national product (GNP). Again in this case the sensitivity of the break date estimate with respect to the lag order is substantial. For $p - 1 = 1$ a break date of 1923 is obtained whereas $p - 1 = 2$ results in $\hat{\eta} = 1933$. Hence, in this case again the model underlying the \mathbf{t} test may not be adequate here for different reasons. Notice, however, that the test values are quite close to the 10% critical value. Hence, even the \mathbf{t} test offers some indication that there may not be a unit root in the series.

The \mathbf{t} test also rejects the unit root in the nominal wage series, in contrast to both other tests. Thus, as in our simulations, the \mathbf{t} test may have power advantages in some situations. Notice, that the τ test value is close to the 10% critical value. The results for the wage series clearly show the virtue of applying both tests simultaneously because it depends on the DGP which test is preferable.

Generally our examples show the virtue of allowing for a break in the DGP when testing for a unit root and they are also in line with the simulation results of the previous section. The break dates vary substantially, however, which is clearly a consequence of the very long observation period covering a number of important events such as two world wars and the Great Crash at the end of the 1920s. For velocity the most important shift seems to have

even occurred in the 19th century. Given our theoretical results the tests should not be sensitive to estimating the break dates poorly. Moreover, the simulation results show that the break date is important in small samples only if the break is quite sizable. Hence, overall the results in Table 1 illustrate some of the asymptotic and small sample properties of our testing procedures.

As mentioned earlier, the unit root properties of similar series were analyzed by a number of other authors as well. For example, Perron (1989), Zivot & Andrews (1992) and Amsler & Lee (1995) have considered the Nelson-Plosser data. It has been argued by Perron (1989) that there may be a break in the slope of the linear trend function in some of the series. Our framework does not cover that type of structural break. Therefore it may be of interest to note that our results are well in line with those of Zivot & Andrews (1992) who also allow for an unknown break date. For instance, for our set of variables, they reject a unit root for real GNP, nominal GNP, real per capita GNP, industrial production, employment and nominal wages (see Zivot & Andrews (1992, Table 6)). Consequently, except for nominal GNP, our tests reject the unit root for all the series as well. In addition we find evidence against a unit root in the CPI and money stock. In the aforementioned study the AR orders are partly larger than in our study. Based on our simulations it seems reasonable to use fairly small lag orders, however, to improve the test power. It may also be worth noting that our estimated break dates differ from those obtained by Zivot & Andrews (1992). Obviously, this does not affect the conclusions from the unit root tests substantially.

7 Conclusions

In this study we have considered tests for unit roots in time series which have a structural level shift at unknown time. The shift is assumed to be representable by a deterministic function. We consider tests which proceed by estimating the shift function and other deterministic terms in a first step, adjust the series for deterministic terms and then apply standard unit root tests to the adjusted series. A test version based on an additive outlier model with known break date is extended to the unknown break date case and a corresponding version for an innovational outlier model is proposed. It is shown that, under our conditions, for any estimate of the break date, the tests have the same asymptotic properties as the

corresponding tests that assume knowledge of the break date.

In a Monte Carlo study the following small sample results are found. Despite the fact that all estimators of the break date lead to the same asymptotic properties of the tests, in small samples it is best to choose the break date which leads to the minimal objective function in the estimation algorithm of the nuisance parameters. The choice of the AR order is of limited importance for the properties of the tests as long as a reasonably large order is used, that is, as long as the order is not understated and not overstated substantially. This result is likely to depend on the sample size and was obtained for a sample of size $T = 100$ in our simulations. A possible strategy for empirical work may be to choose a fairly large AR order in a first round and try also smaller orders if the resulting test does not reject the unit root null hypothesis. Although the asymptotics work for very general shift functions, choosing flexible shift functions may lead to a dramatic loss of power for samples as large as $T = 100$. Therefore it may be best to use a simple shift dummy variable as shift function and apply both types of tests. Depending on the DGP, their relative power may vary substantially, sometimes leading to superior performance of one of the tests and in other situations giving the lead to the other test. Surprisingly, test versions which assume that there is no linear trend in the DGP do not necessarily have sizable power advantages over tests which allow for a linear time trend. Therefore, if there is uncertainty about the possibility of a linear trend, it is preferable to include such a term. Finally, ignoring the break and applying standard DF tests can lead to more or less arbitrary results and is not advisable in practice.

A set of annual U.S. series from the Nelson-Plosser data is used to illustrate the performance of the tests in practice. It turns out that allowing for a break at unknown time, the unit root null hypothesis can be rejected in some cases where rejection at usual significance levels is not possible with standard augmented DF tests. These results indicate that our tests are a useful addition to the toolkit of macroeconometrics.

Appendix. Proof of Theorem 1

In the proof of Theorem 1 we focus on the limiting distribution of the test statistics $\tau(\hat{\eta}, p)$ and $\mathbf{t}(\hat{\eta}, p)$. The limiting distribution of the test statistics $\tau^0(\hat{\eta}, p)$ and $\mathbf{t}^0(\hat{\eta}, p)$ can be obtained

by making straightforward modifications to these proofs.

First consider the test statistic $\tau(\hat{\eta}, p)$. We start by showing that the appropriately standardized moment matrix in the LS estimation of the parameters in (4.1) is asymptotically block diagonal between $\hat{b}_{\hat{\eta}}(L)\Delta G_{t\hat{\eta}}(\hat{\theta}_{\hat{\eta}}) = [\hat{b}_{\hat{\eta}}(L)\Delta f_{t\hat{\eta}}(\hat{\theta}_{\hat{\eta}})' : \hat{b}_{\hat{\eta}}(L)\Delta F_{t\hat{\eta}}(\hat{\theta}_{\hat{\eta}})']'$ and the other regressors. To this end, we first note that (see (A.33) of S&L)

$$\hat{x}_t = x_t - (\hat{\mu}_{\hat{\eta}} - \mu)t - g_{t\hat{\eta}}(\hat{\theta}_{\hat{\eta}})'\hat{\varphi}_{\hat{\eta}} + g_{t\eta}(\theta)'\varphi \quad (A.1)$$

where $\varphi = [\mu_0 : \gamma']'$ and $\hat{\varphi}_{\hat{\eta}} = [\hat{\mu}_{0\hat{\eta}} : \hat{\gamma}'_{\hat{\eta}}]'$. S&L use this identity and their Lemma 3.1 to show that

$$T^{-1/2}\hat{x}_{[Ts]} \xrightarrow{d} \omega G_c(s) \quad (A.2)$$

where $\omega = \sigma/b(1)$. Next conclude from (A.1) that

$$\Delta\hat{x}_t = \Delta x_t - (\hat{\mu}_{\hat{\eta}} - \mu) - \Delta g_{t\hat{\eta}}(\hat{\theta}_{\hat{\eta}})'(\hat{\varphi}_{\hat{\eta}} - \varphi) - \Delta g_{t\hat{\eta}}(\hat{\theta}_{\hat{\eta}})'\varphi + \Delta g_{t\eta}(\theta)'\varphi.$$

Here $\hat{\mu}_{\hat{\eta}} - \mu = O_p(T^{-1/2})$ and $\hat{\varphi}_{\hat{\eta}} - \varphi = o_p(T^a)$ where $\frac{1}{\alpha} < a \leq \frac{1}{4}$ (see Lemma 3.1 of S&L). Thus, because Assumption A(b) implies that $\max_{1 \leq t \leq T} \|g_{t\eta}(\theta)\|$ is bounded uniformly in θ , η and T we have

$$\begin{aligned} \max_{1 \leq t \leq T} |\Delta\hat{x}_t| &= \max_{1 \leq t \leq T} |\Delta x_t| + o_p(T^a) \\ &= o_p(T^a). \end{aligned} \quad (A.3)$$

Here the latter equality can be justified by writing $\Delta x_t = T^{-1}cx_{t-1} + b(L)^{-1}\varepsilon_t$ and using (A.14) of S&L and well-known properties of near integrated processes.

Now write $\hat{w}_t = \hat{b}_{\hat{\eta}}(L)\hat{x}_t = \hat{b}_{\hat{\eta}}(1)\hat{x}_t + \hat{b}_{\hat{\eta}}^*(L)\Delta\hat{x}_t$ and recall from (3.7) of S&L that the coefficients of $\hat{b}_{\hat{\eta}}(L)$ are consistent estimators of the coefficients of $b(L)$. Using these facts in conjunction with (A.2) and (A.3) we readily find that

$$T^{-1/2}\hat{w}_{[Ts]} \xrightarrow{d} \omega G_c(s). \quad (A.4)$$

Because the roots of $\hat{b}_{\hat{\eta}}(L)$ are bounded away from the unit circle by assumption, the coefficients of this polynomial can be bounded by a constant. This fact together with (A.4) and Assumption A(b) shows that

$$\begin{aligned} &\|T^{-1}\sum_{t=p+2}^T \hat{w}_{t-1}[\hat{b}_{\hat{\eta}}(L)\Delta G_{t\hat{\eta}}(\hat{\theta})']\| \\ &\leq \text{const.} \times (T^{-1}\max_{1 \leq t \leq T} |\hat{w}_t|) \sum_{t=1}^T \|\Delta G_{t\hat{\eta}}(\hat{\theta})\| \\ &\leq \text{const.} \times (T^{-1}\max_{1 \leq t \leq T} |\hat{w}_t|) \sup_{\theta \in \Theta, \eta \in N_T} \sum_{t=1}^T \|\Delta G_{t\eta}(\theta)\| \\ &= O_p(T^{-1/2}). \end{aligned} \quad (A.5)$$

Making use of (A.3) instead of (A.4) one can similarly show that, for $1 \leq j \leq p$,

$$\begin{aligned}
& \|T^{-1/2} \sum_{t=p+2}^T \Delta \hat{x}_{t-j} [\hat{b}_{\hat{\eta}}(L) \Delta G_{t\hat{\eta}}(\hat{\theta})']\| \\
& \leq \text{const.} \times (T^{-1/2} \max_{1 \leq t \leq T} |\Delta \hat{x}_t|) \sup_{\theta \in \Theta, \eta \in N_T} \sum_{t=1}^T \|\Delta G_{\eta t}(\theta)\| \\
& = o_p(T^{a-1/2}).
\end{aligned} \tag{A.6}$$

Furthermore, replacing $\Delta \hat{x}_{t-j}$ by unity in these derivations yields

$$\|T^{-1/2} \sum_{t=p+2}^T [\hat{b}_{\hat{\eta}}(L) \Delta G_{t\hat{\eta}}(\hat{\theta})]\| = O_p(T^{-1/2}). \tag{A.7}$$

Combining (A.5) - (A.7) shows that the off-diagonal block of the standardized moment matrix we are considering is of order $o_p(T^{a-1/2}) = o_p(1)$. Using this fact and the inversion formula for a partitioned matrix one can readily see that the same is also true for the inverse of this matrix, provided the smallest eigenvalues of the diagonal blocks are asymptotically bounded away from zero. For the diagonal block corresponding to the regressor $\Delta G_{t\hat{\eta}}(\hat{\theta})$ this follows from Assumption A(c) and for the other diagonal block this becomes clear later in the proof. Thus, since $2a - \frac{1}{2} \leq 0$ and since it will be shown shortly that

$$\sum_{t=p+2}^T [\hat{b}_{\hat{\eta}}(L) \Delta G_{t\hat{\eta}}(\hat{\theta})] r_t = o_p(T^a) \tag{A.8}$$

we can conclude that the asymptotic properties of the LS estimators of the parameters ν , ρ and α_j obtained from the auxiliary regression model (4.1) can be derived by ignoring the regressors $\hat{b}_{\hat{\eta}}(L) \Delta f_{t\hat{\eta}}(\hat{\theta})$ and $\hat{b}_{\hat{\eta}}(L) \Delta F_{t\hat{\eta}}(\hat{\theta})$.

To justify (A.8) we first conclude from (A.1) after straightforward but somewhat tedious algebra that

$$\begin{aligned}
\Delta \hat{w}_t - \frac{c}{T} \hat{w}_{t-1} &= \hat{b}_{\hat{\eta}}(L) \Delta x_t - \frac{c}{T} \hat{b}_{\hat{\eta}}(L) x_{t-1} \\
&\quad - \hat{b}_{\hat{\eta}}(1) (\hat{\mu}_{\hat{\eta}} - \mu) + (\hat{\mu}_{\hat{\eta}} - \mu) \frac{c}{T} \hat{b}_{\hat{\eta}}(L) (t-1) \\
&\quad - \hat{b}_{\hat{\eta}}(L) \Delta f_{t\hat{\eta}}(\hat{\theta}_{\hat{\eta}})' \hat{\gamma}_{\hat{\eta}} + \frac{c}{T} \hat{b}_{\hat{\eta}}(L) g_{t-1, \hat{\eta}}(\hat{\theta}_{\hat{\eta}})' \hat{\varphi}_{\hat{\eta}} \\
&\quad + \hat{b}_{\hat{\eta}}(L) \Delta f_{t\eta}(\theta)' \gamma - \frac{c}{T} \hat{b}_{\hat{\eta}}(L) g_{t-1, \eta}(\theta)' \varphi
\end{aligned}$$

where we can further write

$$\hat{b}_{\hat{\eta}}(L) \Delta x_t = b(L) \Delta x_t + (\hat{b}_{\hat{\eta}}(L) - b(L)) \Delta \hat{x}_t + (\hat{b}_{\hat{\eta}}(L) - b(L)) (\Delta x_t - \Delta \hat{x}_t).$$

Now identify α_j in (4.1) as the j -th coefficient of $\hat{b}_{\hat{\eta}}(L) - b(L)$ and π_1 as $-\hat{\gamma}_{\hat{\eta}} + \gamma$. Then, defining the intercept term ν in an appropriate way, we find from the above that the error

term r_t has the representation

$$\begin{aligned}
r_t = & b(L)\Delta x_t + [\hat{b}_{\hat{\eta}}(L) - b(L)](\Delta x_t - \Delta \hat{x}_t) - \frac{c}{T}\hat{b}_{\hat{\eta}}(L)x_{t-1} \\
& - \hat{b}_{\hat{\eta}}(L)[\Delta f_{t\hat{\eta}}(\hat{\theta}_{\hat{\eta}}) - \Delta f_{t\eta}(\theta)]'\gamma - \hat{b}_{\hat{\eta}}(L)\Delta F_{t\hat{\eta}}(\hat{\theta}_{\hat{\eta}})'\pi_2 \\
& + (\hat{\mu}_{\hat{\eta}} - \mu)\frac{c}{T}\hat{b}_{\hat{\eta}}(L)(t-1) + \frac{c}{T}\hat{b}_{\hat{\eta}}(L)g_{t-1,\hat{\eta}}(\hat{\theta}_{\hat{\eta}})'\hat{\varphi}_{\hat{\eta}} - \frac{c}{T}\hat{b}_{\hat{\eta}}(L)g_{t-1,\eta}(\theta)'\varphi.
\end{aligned} \tag{A.9}$$

The discussion following equation (4.2) of LLS shows that π_2 is a function of $\hat{\theta}_{\hat{\eta}}$ but not a function of other estimators. Thus, $\pi_2 = O_p(1)$ by the compactness of the parameter space Θ and, because the coefficients of $\hat{b}_{\hat{\eta}}(L)$ are consistent estimators of those of $b(L)$, it follows from Assumption A(b) that the contribution of the fourth and fifth terms on the right hand side of (A.9) to the left hand side of (A.8) is of order $O_p(1)$. These arguments and those used to obtain (A.3) also show that the contribution of the first three terms on the right hand side of (A.9) to the left hand side of (A.8) is of order $o_p(T^a)$. Because the sequence $g_{1\eta}(\theta), \dots, g_{T\eta}(\theta)$ is bounded uniformly in θ , η , and T , the same conclusion can be drawn for the last term on the right hand side of (A.9). Finally, the same is also true for the sixth and seventh terms on the right hand side of (A.9) because $\hat{\mu}_{\hat{\eta}} - \mu = O_p(T^{-1/2})$ and $\hat{\varphi}_{\hat{\eta}} = o_p(T^a)$ by Lemma 3.1 of S&L. Thus, we have justified (A.8). In summary, we have shown that the regressors $\hat{b}_{\hat{\eta}}(L)f_{t\hat{\eta}}(\hat{\theta}_{\hat{\eta}})$ and $\hat{b}_{\hat{\eta}}(L)\Delta F_{t\hat{\eta}}(\hat{\theta}_{\hat{\eta}})$ can be ignored when asymptotic properties of the LS estimators of the parameters ν , ρ and α_j in (4.1) are derived.

Next, note that the regressors $\Delta \hat{x}_{t-j}$ ($j = 1, \dots, T$) can also be ignored when asymptotic properties of the LS estimators of ν and ρ are derived. To see this, first note that from (A.1) and the arguments used to obtain (A.3) it is straightforward to conclude that $T^{-3/2} \sum_{t=p+2}^T \hat{w}_{t-1} \Delta \hat{x}_{t-j} = o_p(1)$ whereas $T^{-1} \sum_{t=p+2}^T \Delta \hat{x}_{t-j} = o_p(1)$ ($j = 1, \dots, p$) by (A.2). Thus, the appropriately standardized moment matrix between the regressors $[1 : \hat{w}_{t-1}]'$ and Δx_{t-j} ($j = 1, \dots, p$) is asymptotically block diagonal, and, because the arguments in the proof of Theorem 4.1 of S&L show that the smallest eigenvalues of the two diagonal blocks are asymptotically bounded away from zero, the desired result follows if $T^{-1/2} \sum_{t=p+2}^T \Delta \hat{x}_{t-j} r_t = O_p(1)$ ($j = 1, \dots, p$). This latter fact can be justified by using equation (A.9) and arguments already used above. Thus, we have reduced the problem to essentially the same one as treated in S&L. The main difference is that now an intercept term is included in the auxiliary regression model on which our test is based. This, however, is easily seen to have the same effect on the test as in conventional Dickey-Fuller tests, or, in other words, the resulting limiting distribution becomes the same as its counterpart in LLS

obtained by assuming a known break date. This completes the proof of (4.4) in the case of test statistic $\tau(\hat{\eta}, p)$.

Now consider deriving the limiting distribution of test statistic $\mathbf{t}(\hat{\eta}, p)$. As can be expected from S&L, some of the details are very similar to those in the preceding proof and will therefore be omitted.

In the same way as in the preceding proof we first show that the appropriately standardized moment matrix in the LS estimation of the parameters in (4.2) is asymptotically block diagonal between $\Delta G_{t\hat{\eta}}(\tilde{\theta}_{\hat{\eta}})$ and the other regressors with the off-diagonal elements being at most of order $o_p(T^{1/2-a})$. As far as the regressors \tilde{v}_{t-1} and 1 are concerned, this can be seen by arguments similar to those used for (A.5) and (A.7). So, we only need to consider $\tilde{q}_{t\hat{\eta}}$. From (12.A.10) of Lütkepohl, Müller & Saikkonen (2001) we first conclude that

$$y_t = \mu_* t + k_t + x_t, \quad t = 1, 2, \dots, \quad (\text{A.10})$$

where $\mu_* = \mu/b(1)$ and the sequence k_t has properties similar to $f_{t\eta}(\theta)$ with η and θ fixed. In particular, when Assumption A(b) holds, $\sup_T \sum_{t=1}^T |k_t| < \infty$. From (A.10) one obtains

$$\Delta y_t - \tilde{\mu}_{*\hat{\eta}} = \Delta k_t - (\tilde{\mu}_{*\hat{\eta}} - \mu_*) + \Delta x_t, \quad (\text{A.11})$$

where $\tilde{\mu}_{*\hat{\eta}} - \mu_* = O_p(T^{-1/2})$ by (3.15) and (3.16) of S&L. Thus, the above discussion and arguments used in the case of (A.6) and (A.7) show that the standardized moment matrix between the regressors $\Delta G_{t\hat{\eta}}(\tilde{\theta}_{\hat{\eta}})$ and $\tilde{q}_{t\hat{\eta}}$ is asymptotically block diagonal and the off-diagonal elements are of order $o_p(T^{a-1/2})$. Combining these results we have therefore shown that the standardized moment matrix between $\Delta G_{t\hat{\eta}}(\tilde{\theta}_{\hat{\eta}})$ and all the other regressors in (4.2) is asymptotically block diagonal and the off-diagonal elements are at most of order $o_p(T^{a-1/2})$. An application of the inversion formula for a partitioned matrix then shows that the same is true for the inverse of this matrix, provided the smallest eigenvalues of the diagonal blocks are bounded away from zero. For the diagonal block corresponding to $\Delta G_{t\hat{\eta}}(\tilde{\theta}_{\hat{\eta}})$ this can be seen in the same way as in the preceding proof and for the other diagonal block this will become clear later. Thus, the limiting distribution of the LS estimators of the parameters ν , ρ and π_3 in (4.2) can be derived by ignoring the regressor $\Delta G_{t\hat{\eta}}(\tilde{\theta}_{\hat{\eta}})$ if we have

$$\sum_{t=2}^T \Delta G_{t\hat{\eta}}(\tilde{\theta}_{\hat{\eta}}) e_t^\dagger = o_p(T^a). \quad (\text{A.12})$$

Using equations (4.13), (4.16) and (A.11) of the discussion paper version of LLS to derive an explicit expression for the error term e_t^\dagger in conjunction with (A.11) of the present paper, one can justify (A.12) by arguments similar to those used for (A.8).

Next we demonstrate that the regressor $\tilde{q}_{t\hat{\eta}}$ can also be ignored when asymptotic properties of the LS estimators of ν and ρ in (4.2) are derived. To this end, note that arguments used in the preceding proof in conjunction with the representation of \tilde{v}_t given in the proof of Theorem 1 of Lütkepohl, Müller & Saikkonen (2001) and the consistency properties of the nuisance parameter estimators given in Lemma 3.2 of S&L show that the standardized moment matrix between the regressors $[1 : \tilde{v}_{t-1}]'$ and $\tilde{q}_{t\hat{\eta}}$ is asymptotically block diagonal. Thus, the standardized moment matrix between the three regressors, $[1 : \tilde{v}_{t-1}]'$, $\Delta G_{t\hat{\eta}}(\tilde{\theta}_{\hat{\eta}})$ and $\tilde{q}_{t\hat{\eta}}$ in (4.2), is asymptotically block diagonal. Furthermore, the smallest eigenvalue of each block is asymptotically bounded away from zero. For the first block this will be made clear below, so it suffices to consider the second block. Because $\tilde{\mu}_{\hat{\eta}} - \mu = O_p(T^{-1/2})$ was noticed to hold, the desired conclusion is obtained by repeating the argument given in the proof of Theorem 2 in the discussion paper version of LLS. Thus, applying the inversion formula for a partitioned matrix we can conclude that the limiting distribution of the LS estimators of the parameters ν and ρ can be derived by ignoring also the regressor $\tilde{q}_{t\hat{\eta}}$ provided that

$$\sum_{t=2}^T \tilde{q}_{t\hat{\eta}} e_t^\dagger = O_p(1)$$

holds. However, in the same way as in the case of (A.12), this can be seen by using the explicit expression for the error term e_t^\dagger in conjunction with (A.11) and arguments used in the preceding proof.

Summarizing, the above discussion shows that the limiting distribution of the test statistic $\mathbf{t}(\hat{\eta}, p)$ can be derived by ignoring the regressors $\Delta G_{t\hat{\eta}}(\tilde{\theta}_{\hat{\eta}})$ and $\tilde{q}_{t\hat{\eta}}$, and, in the same way as in the preceding proof, it can further be shown that the limiting distribution of the test statistic $\mathbf{t}(\hat{\eta}, p)$ differs from that of the test statistic \mathcal{T}_2 in S&L only in that it is obtained from an auxiliary model augmented by an intercept term. Thus, the limiting distribution is the same as in the case of a known break date treated in the discussion paper version of LLS. This gives the desired result and completes the proof.

References

- Amsler, C. & J. Lee (1995), An LM test for a unit root in the presence of a structural change, *Econometric Theory*, 11, 359 - 368.
- Banerjee, A., R.L. Lumsdaine & J.H. Stock (1992), Recursive and sequential tests of the unit-root and trend-break hypotheses: Theory and international evidence, *Journal of Business & Economic Statistics*, 10, 271 - 287.
- Christiano, L.J. (1992), Searching for a break in GNP, *Journal of Business & Economic Statistics*, 10, 237 - 250.
- Evans, G.W. (1989), Output and unemployment dynamics in the United States: 1950 - 1985, *Journal of Applied Econometrics*, 4, 213 - 237.
- Fuller, W.A. (1976), *Introduction to Statistical Time Series*, New York: Wiley.
- Kleibergen, F. & H. Hoek (1999), Bayesian analysis of ARMA models using noninformative priors, Discussion Paper, Econometric Institute and Tinbergen Institute, Erasmus University Rotterdam.
- Lanne, M., H. Lütkepohl & P. Saikkonen (2001), Comparison of unit root tests for time series with level shifts, *Journal of Time Series Analysis*, forthcoming, Discussion Paper No. 88, 1999, SFB 373, Humboldt-Universität zu Berlin, <http://sfb.wiwi.hu-berlin.de>.
- Leybourne, S., P. Newbold & D. Vougas (1998), Unit roots and smooth transitions, *Journal of Time Series Analysis*, 19, 83 - 97.
- Lütkepohl, H. (1991), *Introduction to Multiple Time Series Analysis*, Berlin: Springer-Verlag.
- Lütkepohl, H., C. Müller & P. Saikkonen (2001), Unit root tests for time series with a structural break when the break point is known, Chapter 12 in C. Hsiao, K. Morimune & J. Powell (eds.), *Nonlinear Statistical Modeling: Essays in Honor of Takeshi Amemiya*, Cambridge: Cambridge University Press, pp. 327 - 348.
- Nelson, C.R. & C.I. Plosser (1982), Trends and random walks in macroeconomic time series: Some evidence and implications, *Journal of Monetary Economics*, 10, 139 - 162.

- Perron, P. (1989), The great crash, the oil price shock and the unit root hypothesis, *Econometrica*, 57, 1361 - 1401.
- Perron, P. (1990), Testing for a unit root in a time series with a changing mean, *Journal of Business & Economic Statistics*, 8, 153 - 162.
- Perron, P. & T.J. Vogelsang (1992), Nonstationarity and level shifts with an application to purchasing power parity, *Journal of Business & Economic Statistics*, 10, 301 - 320.
- Saikkonen, P. & H. Lütkepohl (2001a), Testing for unit roots in time series with level shifts, *Allgemeines Statistisches Archiv*, 85, 1 - 25.
- Saikkonen, P. & H. Lütkepohl (2001b), Testing for a unit root in a time series with a level shift at unknown time, *Econometric Theory*, forthcoming, Discussion Paper No. 72, 1999, SFB 373, Humboldt-Universität zu Berlin, <http://sfb.wiwi.hu-berlin.de>.
- Zivot E. & D.W.K. Andrews (1992), Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis, *Journal of Business & Economic Statistics*, 10, 251 - 270.

Figure 1: Frequency distributions of estimators for the break date for DGP (5.1), $T = 100$, $b_1 = 0.5$, $p = 2$, true break date $\tau = 50$ and different values of γ .

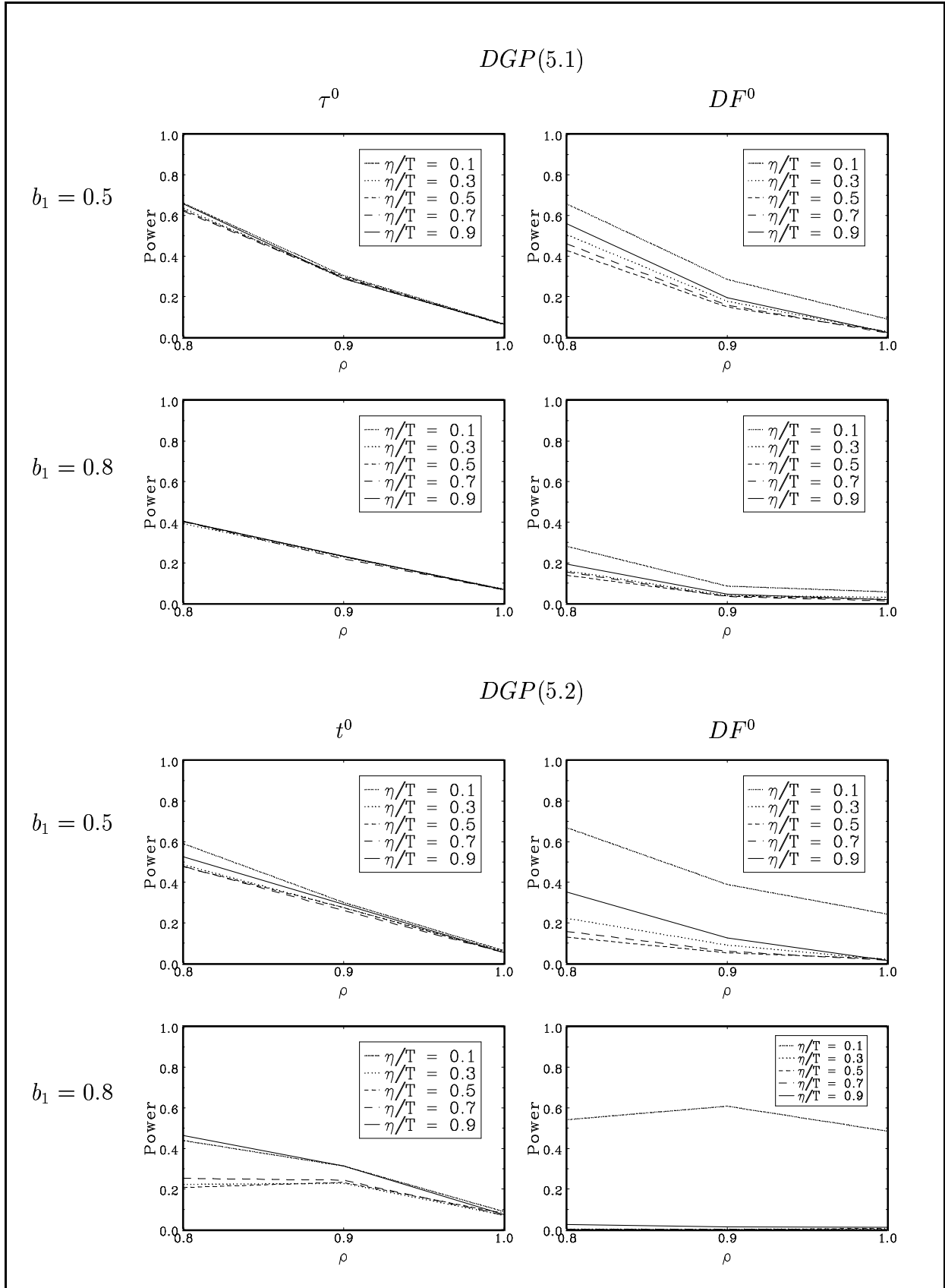


Figure 2: Impact of break date on the tests without linear trend term for $T = 100$ and $\gamma = 1$.

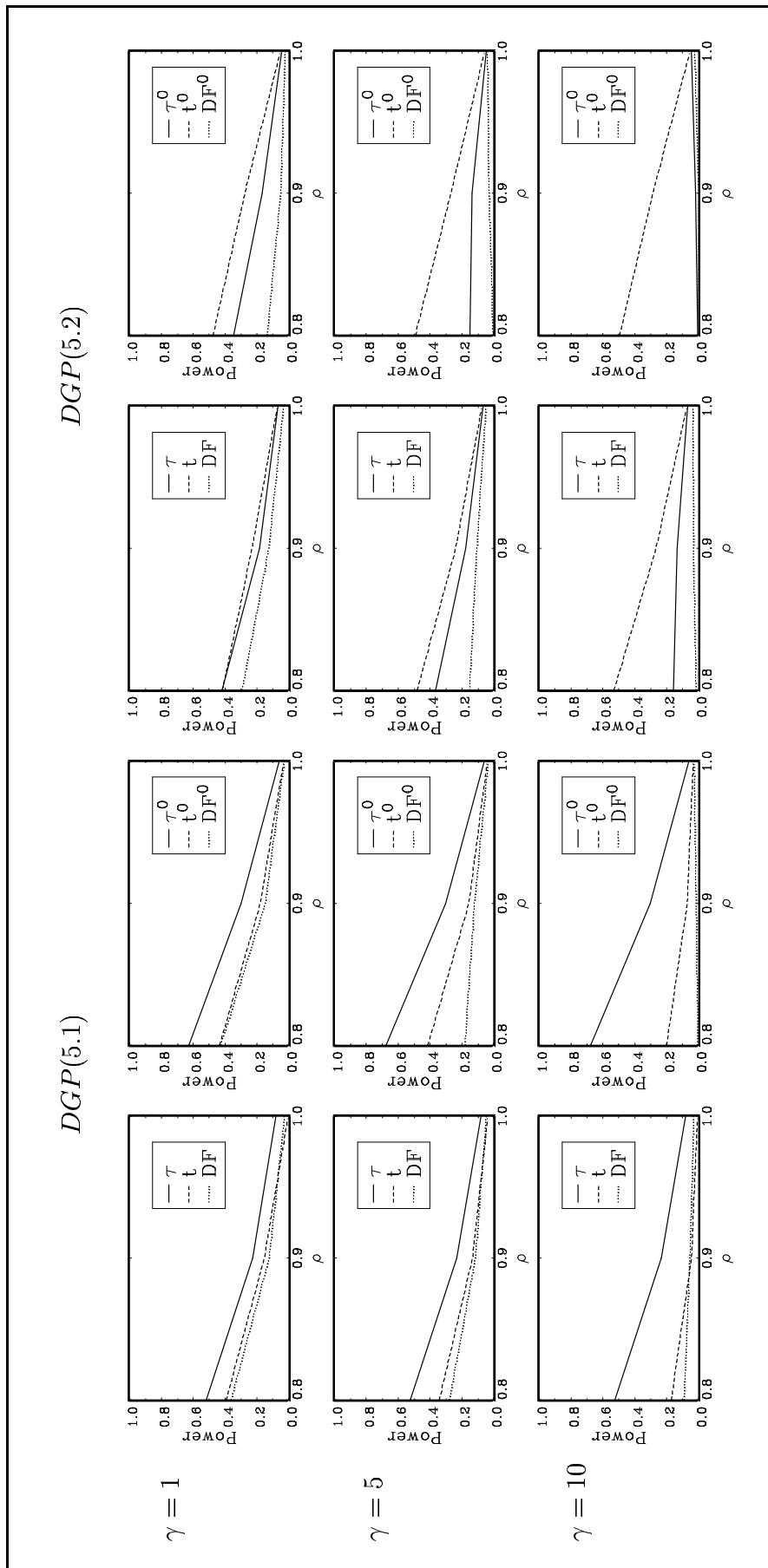


Figure 3: Impact of the size of the break on the tests for $T = 100$, $b_1 = 0.5$.

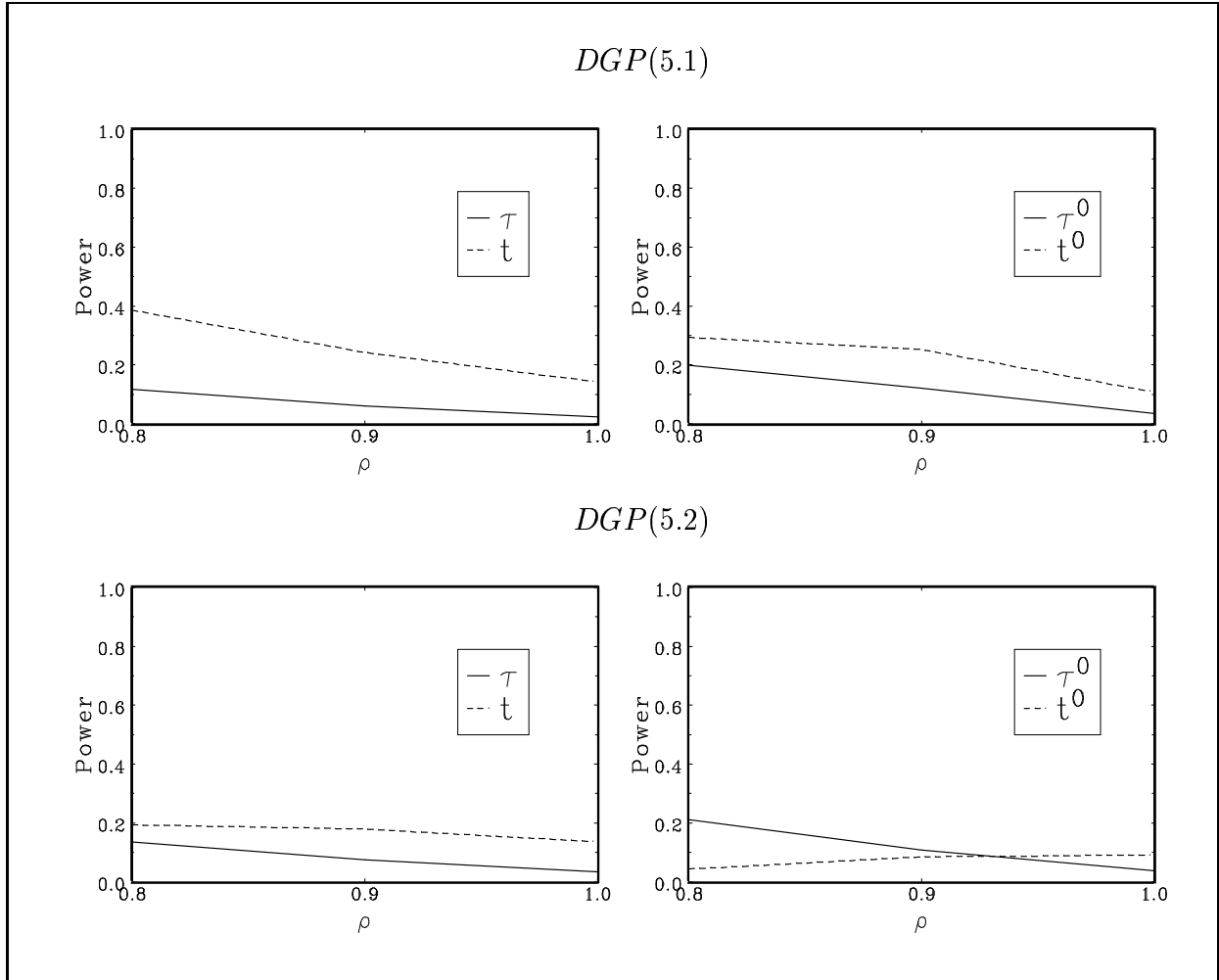


Figure 4: Power functions of the tests with a smooth shift function for $T = 100$, $\gamma = 5$ and $b_1 = 0.5$.

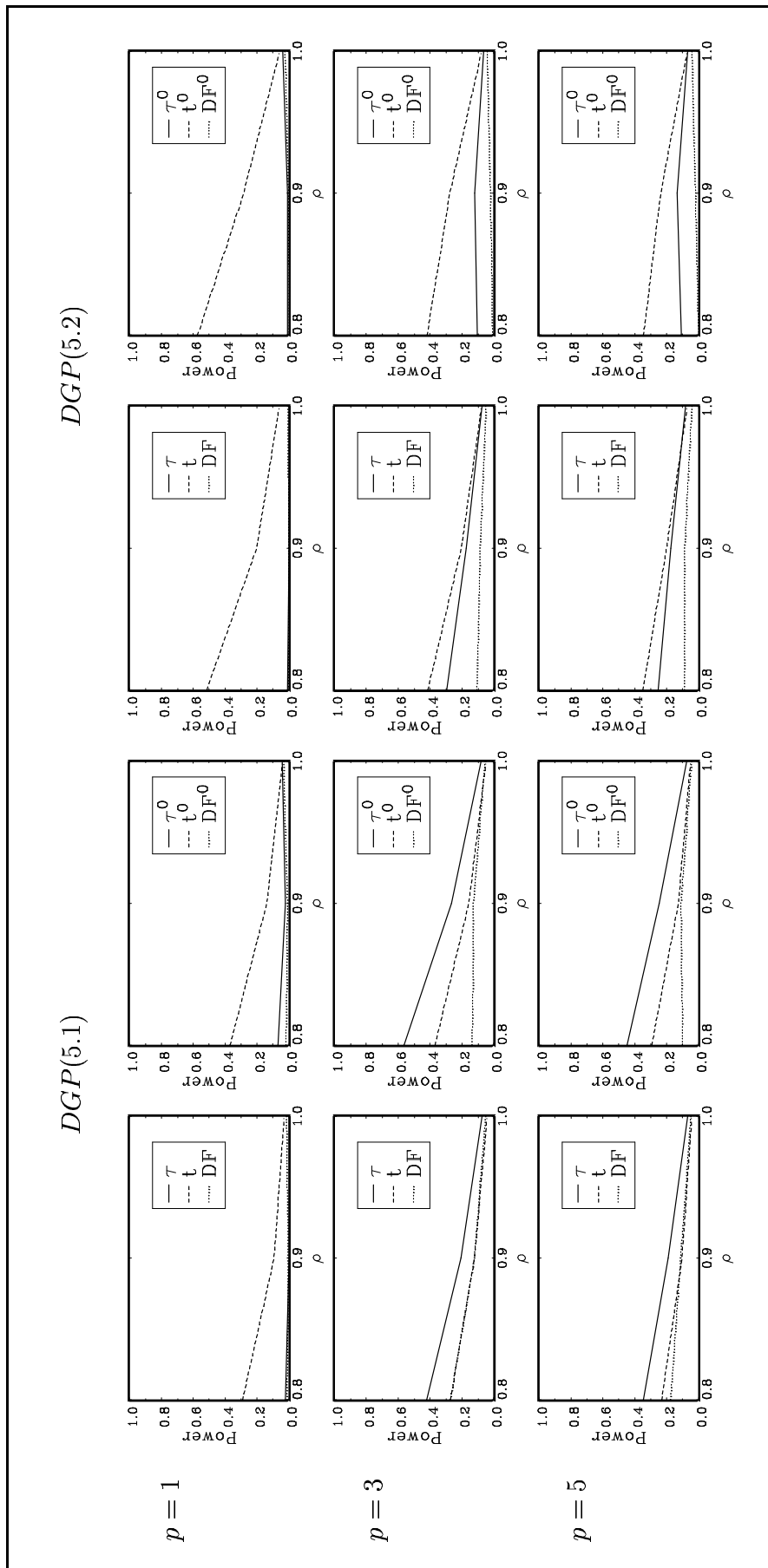


Figure 5: Impact of AR order on the tests for $T = 100$, $\gamma = 5$ and $b_1 = 0.5$.